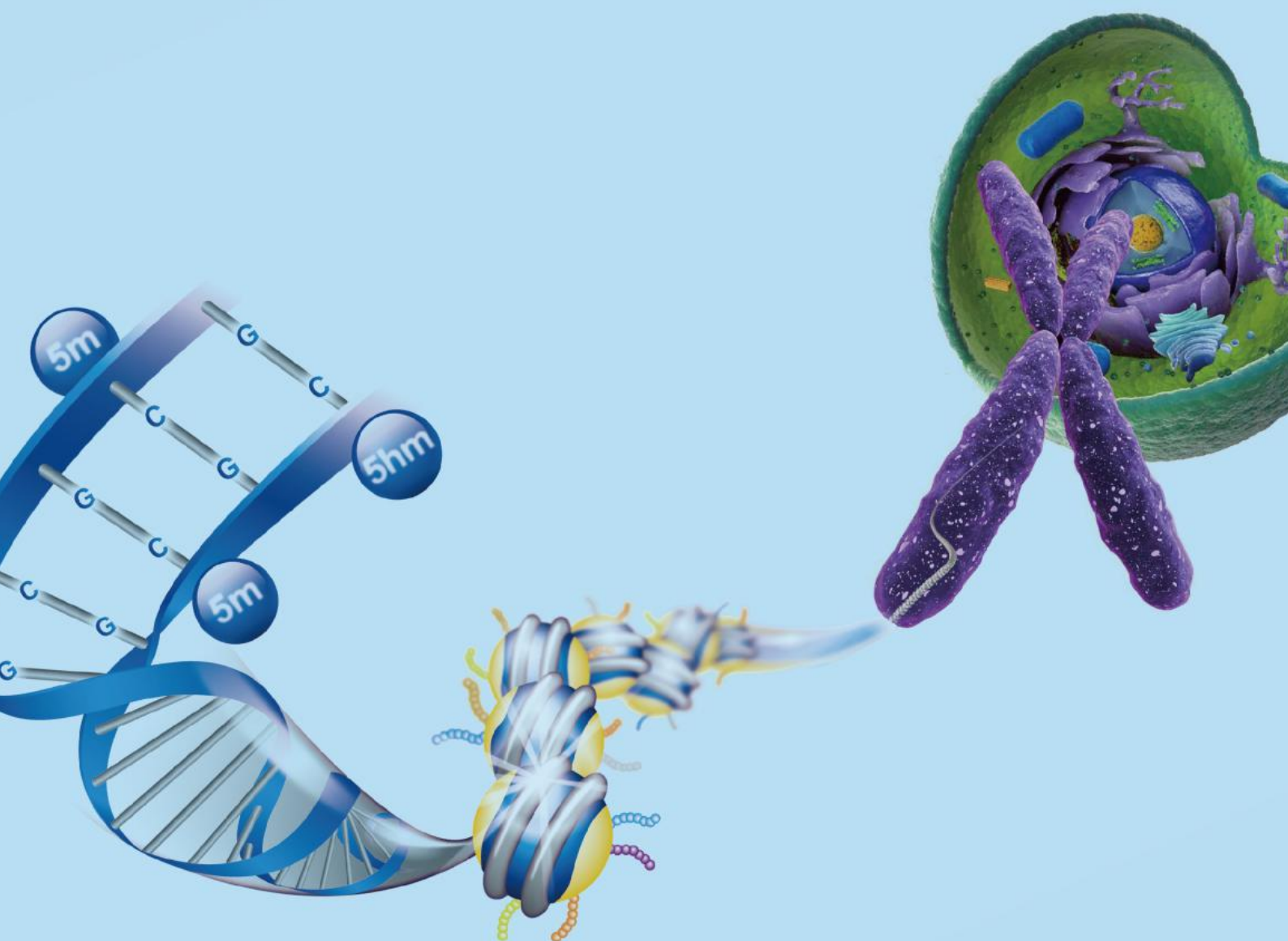


伯豪生物全新推出甲基化芯片服务

Illumina Infinium Methylation EPIC BeadChip

——为您奉上前沿的甲基化研究方案



上海伯豪生物技术有限公司

SHANGHAI BIOTECHNOLOGY CORPORATION

咨询热线: 800-820-5086/400-880-5086

电话: 021-58955370

网址: www.shbio.com

邮箱: market@shbio.com

地址: 上海张江高科技园李冰路151号

芯片简介

Illumina在原450K芯片巨大成功的基础上，推出了新一代的DNA甲基化芯片Illumina Infinium Methylation EPIC BeadChip芯片（下文中以850K芯片代称），为研究者提供了一个可靠且经济高效的甲基化分析平台。850K芯片不但保持了对CpG岛，基因启动子区的全面覆盖，还特别加强了增强子区（新增了333,265个探针覆盖来自ENCODE及FANTOM5计划的增强子）以及基因编码区的探针覆盖。广泛应用于干细胞研究、肿瘤和其他复杂疾病研究，是目前最适合表现基因组全关联分析研究的全基因组DNA甲基化芯片。

芯片特点

- 覆盖全面：检测853,307个CpG位点；全面覆盖CpG岛、启动子、编码区及增强子。
- 无需“甲基化DNA免疫共沉淀”，亚硫酸氢盐处理基因组DNA即可进行芯片实验；Infinium探针设计，直接识别甲基化位点。
- 技术重复重现率>98%。
- 操作简单，无须PCR扩增；模板量可低至250ng。
- 适用于FFPE样品。

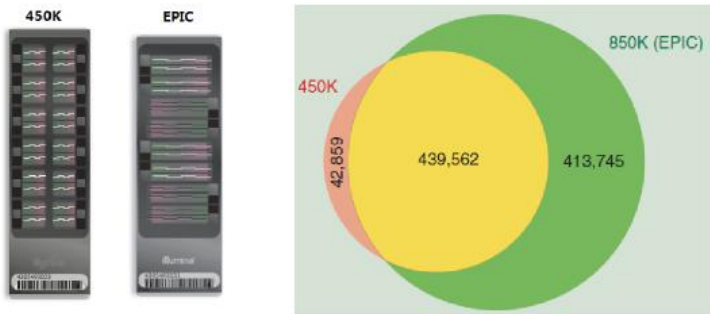


图 1 Illumina Infinium MethylationEPIC (850K)芯片与450K芯片的异同。

芯片对全基因组全覆盖

850K芯片对基因组的功能元件实现了空前全面覆盖,下图及表格统计展示了850K芯片对于不同染色体、基因组功能元件的覆盖及分布。以上的设计不但可从泛-增强子及编码区域角度分析甲基化组，还在全基因组表现遗传关联研究（EWAS）中发挥重要作用。

850K芯片中包含的一系列有价值的位点：

- CpG岛以外的CpG位点
- 人类肝细胞中的非CpG甲基化位点（CHH位点）
- 肿瘤（多种类型的癌症）VS正常样本中的不同甲基化位点，可以检测多种样本类型
- FANTOM5增强子
- ENCODE开放染色质和增强子
- 脱氧核糖核酸酶超敏位点
- miRNA启动子区域
- 涵盖Infinium Methylation 450 BeadChip芯片中的 > 90%的内容

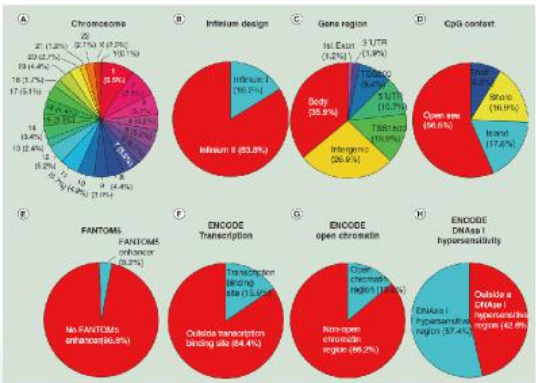


图 2 850K芯片探针在不同染色体及基因组功能元件中的分布。

探针分布比例	探针分布区域
74.40%	编码基因
0.85%	非编码RNA
24.60%	不与转录本关联

探针分布比例	探针分布区域
30.90%	CpG岛
23%	CpG shores
9.70%	CpG shelves
36.30%	open sea

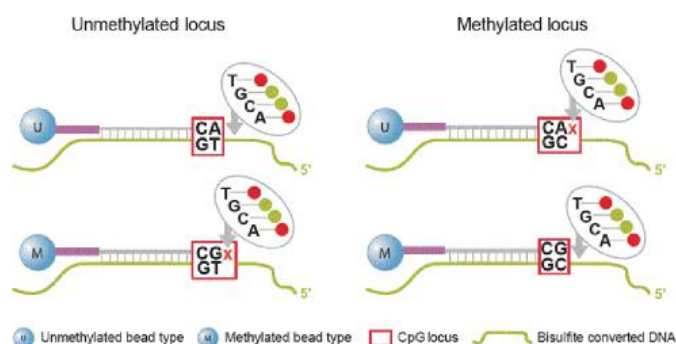
探针分布比例	探针分布区域
41%	promoter(TSS, exon1,5'UTR)
3.20%	3'UTR
30.90%	gene body
24.60%	intergenic

芯片数据的可靠性

850K芯片沿用了在450K芯片中取得成功的Infinium I及II探针设计,以达到保证数据稳定可靠的同时也达到对目标区域全面覆盖。

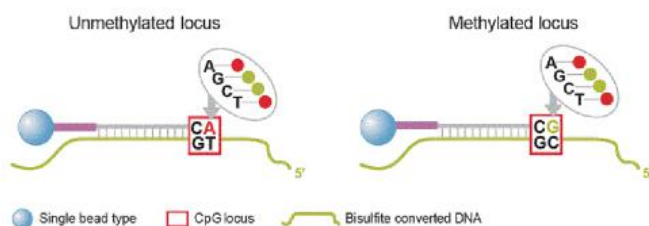
探针类型	探针数/位点数	反应类型	延伸所用的核苷酸	区分方式
Infinium I 探针	两根探针检测一个位点	延伸反应	dNTP-BioT	看信号的有无
Infinium II 探针	一根探针检测一个位点	单碱基延伸	ddNTP-BioT ddNTP-DNP	荧光颜色区分

(1) Infinium I 探针设计



在InfiniumI设计中对于每个甲基化位点,都对应设计有两种探针:M型磁珠、U型磁珠。M型磁珠尾部为G,用来检测甲基化位点(C)。U型磁珠尾部为A,用来检测未甲基化位点(T)。根据单碱基延伸的原理,仅当探针最后一个碱基与模板配对时,荧光标记的核苷酸才能掺入并被检测到荧光信号。通过分析M型与U型磁珠荧光信号的结果计算甲基化值。

(2) Infinium II 探针设计



Infinium II 探针只使用一种磁珠,探针末端为C,与目的位点的前一个碱基配对,只延伸一个碱基(ddNTP-BioT, ddNTP-DNP)分别与非甲基化或甲基化位点配对。根据两种荧光类型计算目的位点的甲基化程度。

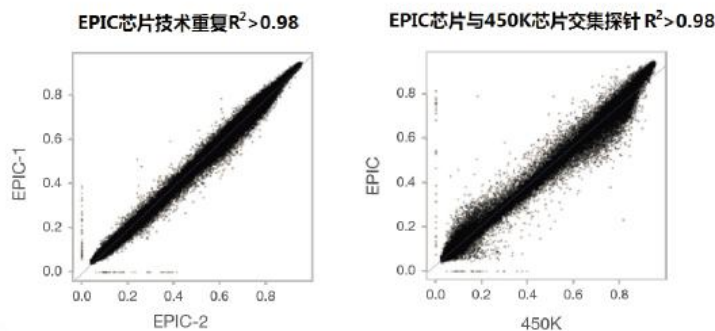


图3 EPIC芯片自身技术重复间的相关性以及450K相同探针间的相关性R方都在0.98以上。

适用于FFPE样品的甲基化研究

Illumina为FFPE样本的检测改进了protocol，以获得更可靠稳定的结果。

Table 1: Comparative Infinium Methylation Data Quality Metrics-Standard vs. FFPE

Methylation BeadChip	Standard Protocol	FFPE Protocol
Reproducibility (Technical Replicates)	$r^2 \geq 98\%$	$r^2 \geq 98\%$
Number of sites detected*	$\geq 96\%$	$\geq 90\%$

*Based on noncancer samples, recommended sample input amounts of high-quality DNA as confirmed by PicoGreen and following all other Illumina recommendations as per User Guides.

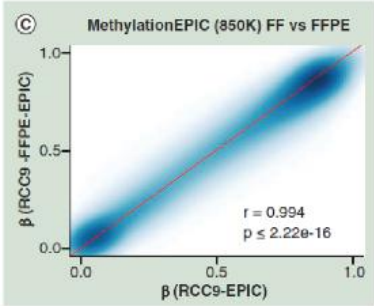


图 4 850K芯片用于FFPE样本时的技术重复相关性（R²>0.98）及检出率(>90%); 相同标本分别用FFPE处理与冰冻效果相比，重复相关性（R=0.994）。

样本要求

- 样品类型：组织、细胞、基因组DNA
- 1.样品纯度：OD 260/280值应在1.7 ~ 2.0 之间；RNA 应该去除干净；不得有其它个体或其它物种的DNA污染。
 - 2.样品浓度：浓度不低于50ng/μl。
 - 3.样品总量：每个样品总量不少于500ng。
 - 4.样品溶剂：溶解在TE中。
 - 5.样品运输：DNA低温运输（-20℃）；在运输过程中请用parafilm将管口密封好，以防出现污染。

数据分析

基础分析	高级分析
1. 原始数据预处理	1. 组内样本甲基化程度展示
2. 样本的beta值密度曲线	2. T-sne降维算法展示样本分组情况
3. 数据归一化（甲基化位点总表，含注释信息）	3. 甲基化与基因表达谱的关联分析
4. 差异甲基化位点（含注释信息）	4. 特定基因甲基化与基因表达关联分析
5. 差异甲基化位点的可视化展示	5. 组间启动子区甲基化差异分析
6. 按基因注释分类得到差异甲基化区域(6类)	6. 特定基因甲基化程度走势分析
7. 按CpG注释分类得到差异甲基化区域(5类)	7. 差异甲基化位点染色体展示（圈图）
	8. 所有基因甲基化程度走势分析
	9. CpG岛上下游2K甲基化程度走势分析
	10. 不同类型基因甲基化程度走势分析
	11. miRNA转录起始位点上下游5K甲基化程度走势分析
	12. DMR分析（Bumphunter）
	13. 基于甲基化数据的肿瘤亚型分析
	14. EWAS分析
	15. 基于甲基化芯片进行CNV分析
	16. 基于甲基化芯片的克隆演化分析
	17. 其它高级分析内容。（需要定制）

分析结果展示

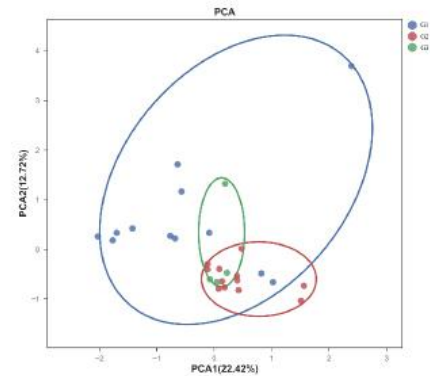


图5 采用主成分统计方法计算样本之间的距离，图中的点（对象）的距离反应了它们的相似性或差异性（不相似性）。

Gene/CpG island probe distribution

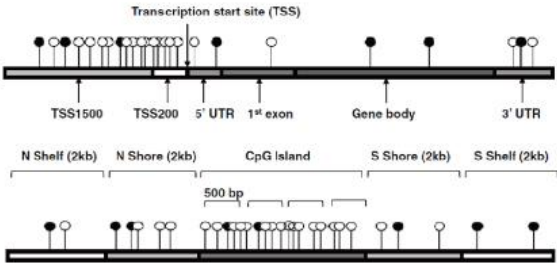


图6 按基因注释分类(TSS1500,TSS200,5' UTR,1stexon,Gene body,3' UTR 6种区域)和CpG岛注释分类（N Shelf, N Shore, CpG Island, S Shore, S Shelf等5种区域）分别筛选样本组间的差异甲基化位点和区域。

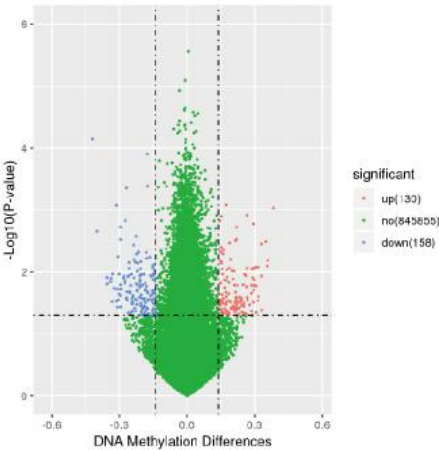


图7 火山图展示两组间的所有差异甲基化位点的结果，红色的点为实验组相对显著高甲基化位点，蓝色的点为实验组相对显著低甲基化位点。

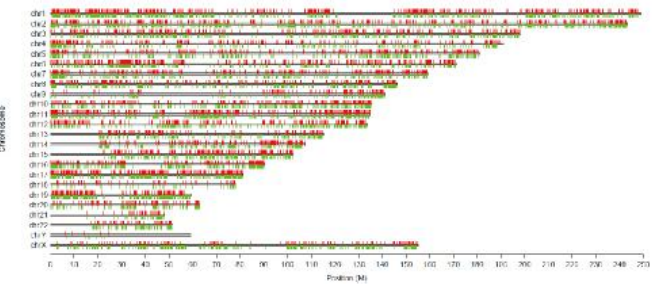


图8 差异甲基化位点在染色体定位，高甲基化或低甲基化的基因是否有染色体的偏向性或成簇分布的特点。其中红色代表实验组相对高甲基化位点，绿色代表实验组相对低甲基化基因位点。

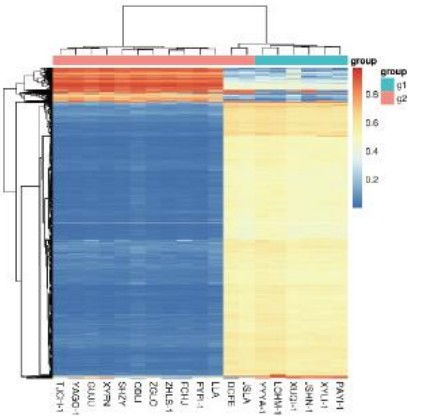


图9 聚类热图展示显著差异的CpG甲基化位点在不同组间各样本中的分布，并可作为样本分类的标志物。

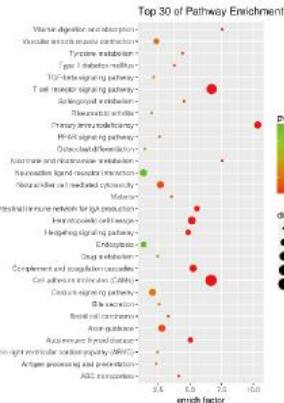


图10 取显著差异位点对应的基因进行GO/KEGG富集分析，enrich factor代表富集程度，pvalue指富集显著与否。

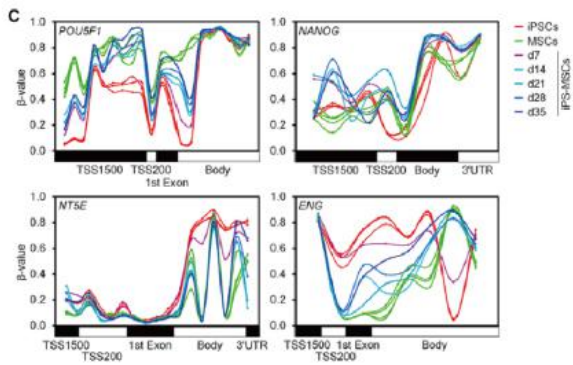


图11 特定基因的甲基化程度走势图，可直观展示不同样本在同一基因不同基因功能区域的甲基化程度变化。

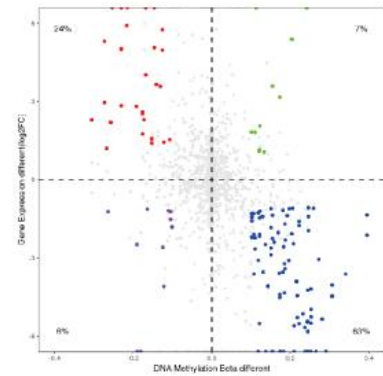


图12 850K甲基化与基因表达谱的关联分析，协助筛选DNA甲基化与基因表达显著正相关或负相关的基因。

甲基化芯片研究案例

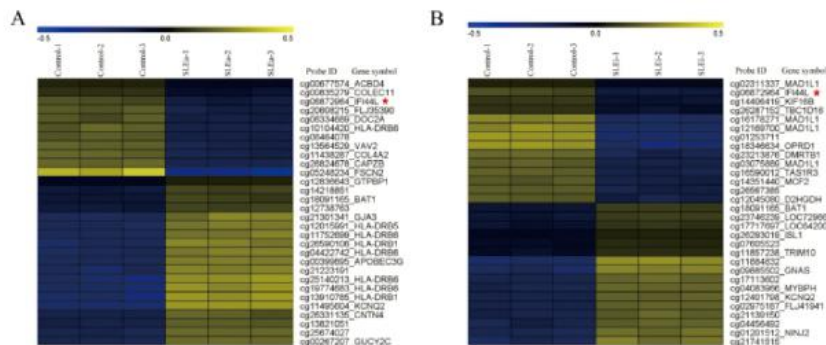
利用甲基化芯片及焦磷酸测序发现系统性红斑狼疮新型诊断标志物

Zhao M, Zhou Y, Zhu B, et al.. IFI44L promoter methylation as a blood biomarker for systemic lupus erythematosus. Ann Rheum Dis. 19 January 2016 [PMID: 26787370]

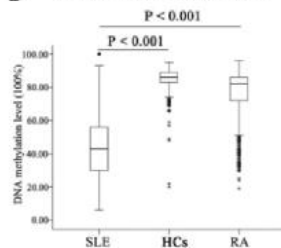
该项研究发现IFI44L基因甲基化水平能够区分系统性红斑狼疮患者与正常人及其他自身免疫性疾病，显著提高了系统性红斑狼疮的诊断可靠性和准确度（特异性95%以上，敏感性90%以上），而且还可应用于系统性红斑狼疮的疗效判断。该研究中Illumina 450K芯片服务由伯豪生物提供。

选择IFI44L作为候选标志物的原因：

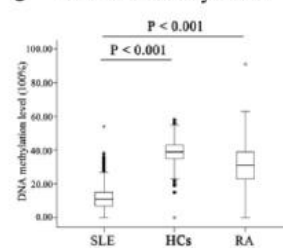
1. IFI44L的所有显著差异CpG位点均位于转录起始位点上游1500bp范围和5' -UTR区域内。
2. IFI44L_cg06872964检测到的CpG位点在活动期患者和稳定期患者外周血DNA中甲基化水平都显著低于健康对照。
3. 课题组之前研究结果也曾发现，在红斑狼疮患者的CD4+细胞和中性粒细胞中IFI44L基因启动子都呈现低甲基化状态。



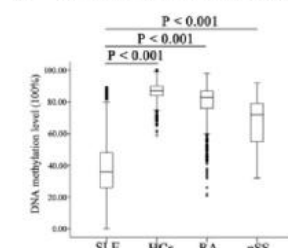
B Site1 in discovery cohort



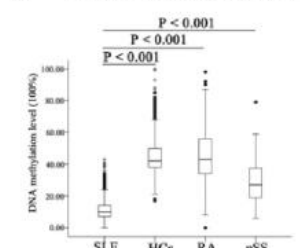
C Site2 in discovery cohort



A Site1 in validation cohort 1



B Site2 in validation cohort



在红斑狼疮患者的外周血单核细胞DNA中，Site 1和Site 2这两个CpG位点的甲基化水平明显低于健康对照和类风湿性关节炎患者。

扩大样本验证中国患者对候选标志物的特异性和敏感性。